

Weekly Report

Lu Junhua

2016 年 7 月 25 日

Done

- Finish Chapter 6 and the beginning of Chapter 7 of *DOM Scripting: Web design with JavaScript and the Document Object Model*.
- Revise the some parts of vis paper.

I discussed with Cai Xiwen and write lists of tasks and post-task questions, to enable a task-based expert evaluation.

I revised all the part related to model. Prof. Peng provides suggestions, and helped revise what I have written. I also Wrote a cover letter which describes what we have done for revising model-related problems.

We discuss the new idea of design, and design alternatives of two visual encodings.

To do

- We will go to netease to have an interview with the experts.
- If the system is finished, the major part of revision should be done.

论文

- WWW16 *Visualizing large-scale and high-dimensional data*: 一开始以为是一篇可视化文章, 后来发现仍然是一篇算法类型的, 文章标题取的比较大. 文章提出了一种比t-SNE更好的降维可视化方法, 降到(2-3维), 可以对网络型, 文本型, 图像数据等进行处理, 取得较好效果, 在可视化手段上没有特别贡献, 但是展现的结果可以看出确实优于t-SNE, 且时间更快.

计算部分, 是快速并且精确的逼近K-临近图. 这个概念我第一次是在我们课程论文中看到, 作为流形降维的一个保形的手段. 而在这里, 作者利用random projection trees的方法, 结合“我邻居的邻居很大可能还是我的邻居”的思想, 通过迭代快速构建了临近图, 并能达到很高的精度(接近百分之百). 图中边的权重采用的是和t-SNE一样的方法. 后面在可视化的时候, 则是利用principled probabilistic model构建一个随机图模型, 将高维点与低位空间中的点作映射, 通过优化一个极大似然函数来解出低维中点的距离, 进而作出位置.

在可视化上比较百万数据集, 本文方法结果看起来比t-SNE更为自然. 效率, accuracy上本文方法也远优于t-SNE. 在非线性降维可视化中, 这种方法有可能称为新潮流.